ECON 241B Midterm

1) Consider the following linear predictor equation:

$$Recid_{t+1} = \alpha + Recid_t\beta + u$$

where $Recid_t$ is the recidivism rate in time t.

(a) Show that:

$$\alpha = \mu_{Recid_{t+1}} - \mu_{Recid_t}\beta$$

and

$$\beta = \frac{Cov(Recid_t, Recid_{t+1})}{Var(Recid_t)}$$

First, note that:

$$\mathbb{E}[Recid_{t+1}] = \mathbb{E}[\alpha] + \mathbb{E}[Recid_t\beta] + \mathbb{E}[u]$$

Which simplifies to:

$$\mu_{Recid_{t+1}} = \alpha + \mu_{Recid_t}\beta + 0$$

Which can be rearranged to give us α. Turning to β, first note that if we subtract α from both sides of our equation we get:

$$Recid_{t+1} - \mu_{Recid_{t+1}} = (Recid_t - \mu_{Recid_t})\beta + u$$

From here we can follow the normal steps for calculating β, giving us:

$$\left(\mathbb{E}[(Recid_t - \mu_{Recid_t})(Recid_t - \mu_{Recid_t})]\right)^{-1} \mathbb{E}[(Recid_t - \mu_{Recid_t})(Recid_{t+1} - \mu_{Recid_{t+1}})]$$

which of course is equal to:

$$\beta = \frac{Cov(Recid_t, Recid_{t+1})}{Var(Recid_t)}$$

(b) Now, suppose we are willing to assume that $\mu_{Recid_t} = \mu_{Recid_{t+1}} = \mu$. Re-write the conditional expectation as a function of μ.

$$\mathbb{E}[Recid_{t+1}|Recid_t] = \mathbb{E}[\alpha|Recid_t] + \mathbb{E}[Recid_t|Recid_t]\beta + \mathbb{E}[u|Recid_t]$$

Which can be simplified to:

$$\mathbb{E}[Recid_{t+1}|Recid_t] = (1-\beta)\mu + Recid_t\beta + \mathbb{E}[u|Recid_t]$$

(c) Assume now that $Var(Recid_t) = Var(Recid_{t+1})$. Prove that $\beta \leq 1$. Does this imply that recidivism rates are converging? Why or why not?

Note that assuming the variances are equal means that we can write:

$$\beta = Corr(Recid_{t+1}, Recid_t)$$

Which we know is between -1 and 1. This does not mean that recidivism rates are converging as we derived $\beta \leq 1$ from the assumption that $Var(Recid_t) = Var(Recid_{t+1})$. This exercise has been an example of the regression fallacy.

2) Estimating the error variance. Recall the "natural" estimator of the error variance $\tilde{\sigma}^2 = \frac{1}{n}u^T u$ Hint: $M = I_n - P$

(a) Why do we not use the "natural" estimator? What do we use instead?

We don't observe u. $\hat{\sigma}^2 = \frac{1}{n}\hat{u}^T\hat{u}$

(b) Write $\hat{\sigma}^2$ as a function of u

$$\begin{aligned}
\hat{\sigma}^2 &= n^{-1}\hat{u}^T\hat{u} \\
&= n^{-1}y^T MMy && (My = \hat{u}) \\
&= n^{-1}y^T My && \text{(M is Idempotent)} \\
&= n^{-1}u^T Mu && (My = Mu)
\end{aligned}$$

(c) Prove that $\tilde{\sigma}^2 - \hat{\sigma}^2 \geq 0$. What does this imply about our variance estimator?

$$\begin{aligned}
\tilde{\sigma}^2 - \hat{\sigma}^2 &= n^{-1}u^T u - n^{-1}\hat{u}^T\hat{u} \\
&= n^{-1}u^T u - n^{-1}u^T Mu && \text{(Shown above)} \\
&= n^{-1}u^T(I_n - M)u \\
&= n^{-1}u^T Pu \\
&\geq 0 && \text{(P is positive-semidefinite, term is a quadratic form)}
\end{aligned}$$

Therefore we know that our estimated variance will be biased towards zero.

3) Suppose we are interested in learning about the relationship between work experience and wages, so we estimate the following linear projection:

$$\mathcal{P}(log(wage)|exp) = 1 + .5exp$$

(a) Is our projection an exact, or approximate, expression for the mean of the wage distribution conditional on experience?

Proposed Solution:

Although experience is a discrete variable, given that it takes a large number of values we are approximating the conditional mean.

(b) What assumptions are required for us to assume that our estimate is causal? Do you think they are likely to hold in this example? Why or why not?

Proposed Solution:

For our estimate to be interpreted causally we need both:

$$E[y|x] = x^T\beta$$

and

$$\Delta_x e = 0$$

This is unlikely to be true in our estimate for a variety of reasons, i.e. both assumptions likely fail. Accepted answers may discuss non-linearity of the CEF, omitted variables, etc.